JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

Commonly Applied Structural Models with Latent Variables

Statistics for Psychosocial Research II:
Structural Models

Qian-Li Xue

Outline

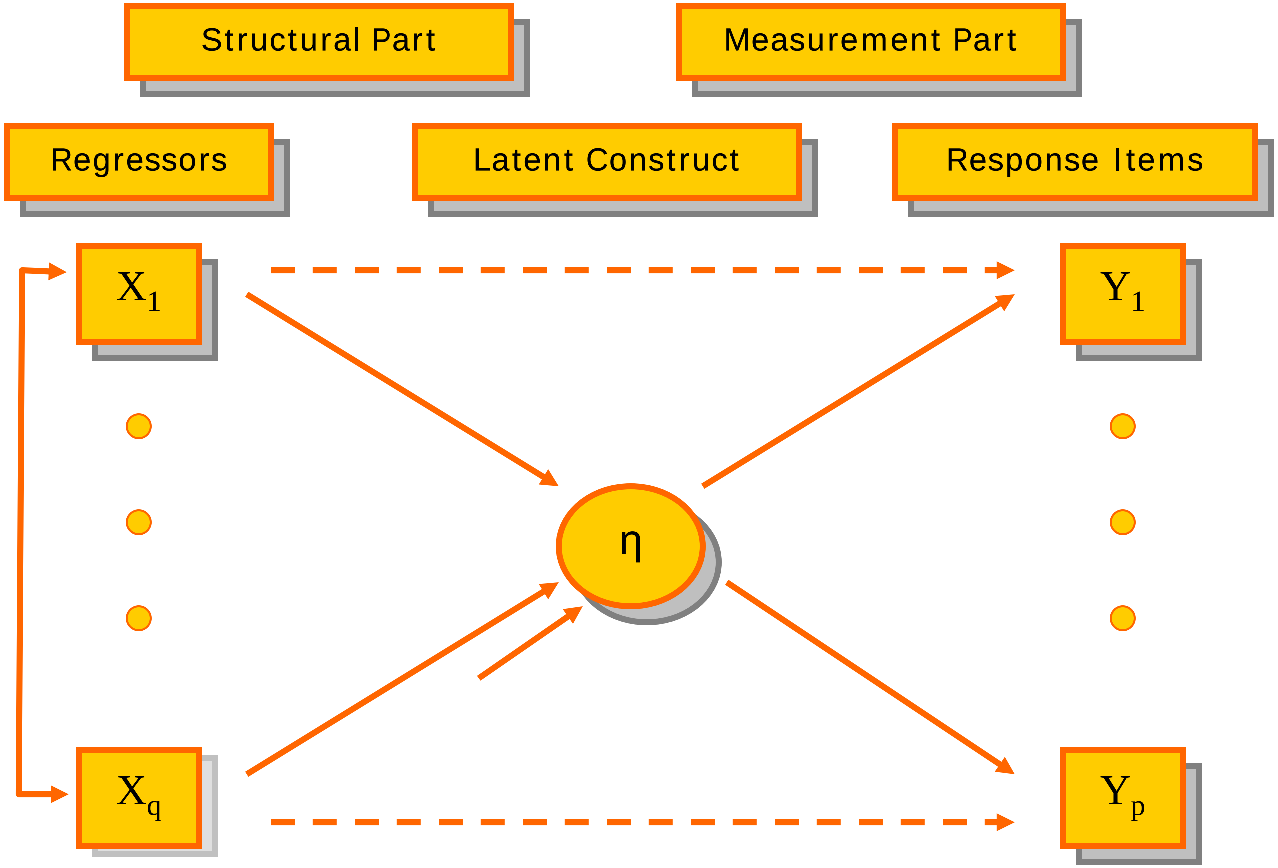
1. Three useful types of SEMs
   a) Multiple Indicators Multiple Causes (MIMIC) model
   b) Multitrait-Multimethod SEM
   c) Causal vs. Effect Indicators

2. Group comparison models

1a) Multiple Indicators Multiple Causes (MIMIC) Model

The MIMIC Structural Equation Model

(Jöreskog & Goldberger, 1975)



Adapted from Muthén (1988)

A MIMIC Structural Probit Model

(Muthén, 1988, 1989)

- Extension of standard Item Response Theory (IRT) modeling of dichotomous items to include covariates that
- Simultaneously addresses four issues:
  - Estimation of IRT measurement parameters (internal validity)
  - Assessment of associations between covariates and the latent trait (external validity)
  - Detection of item bias (differential item performance)
  - Relaxation of the assumptions of unidimensionality and conditional independence

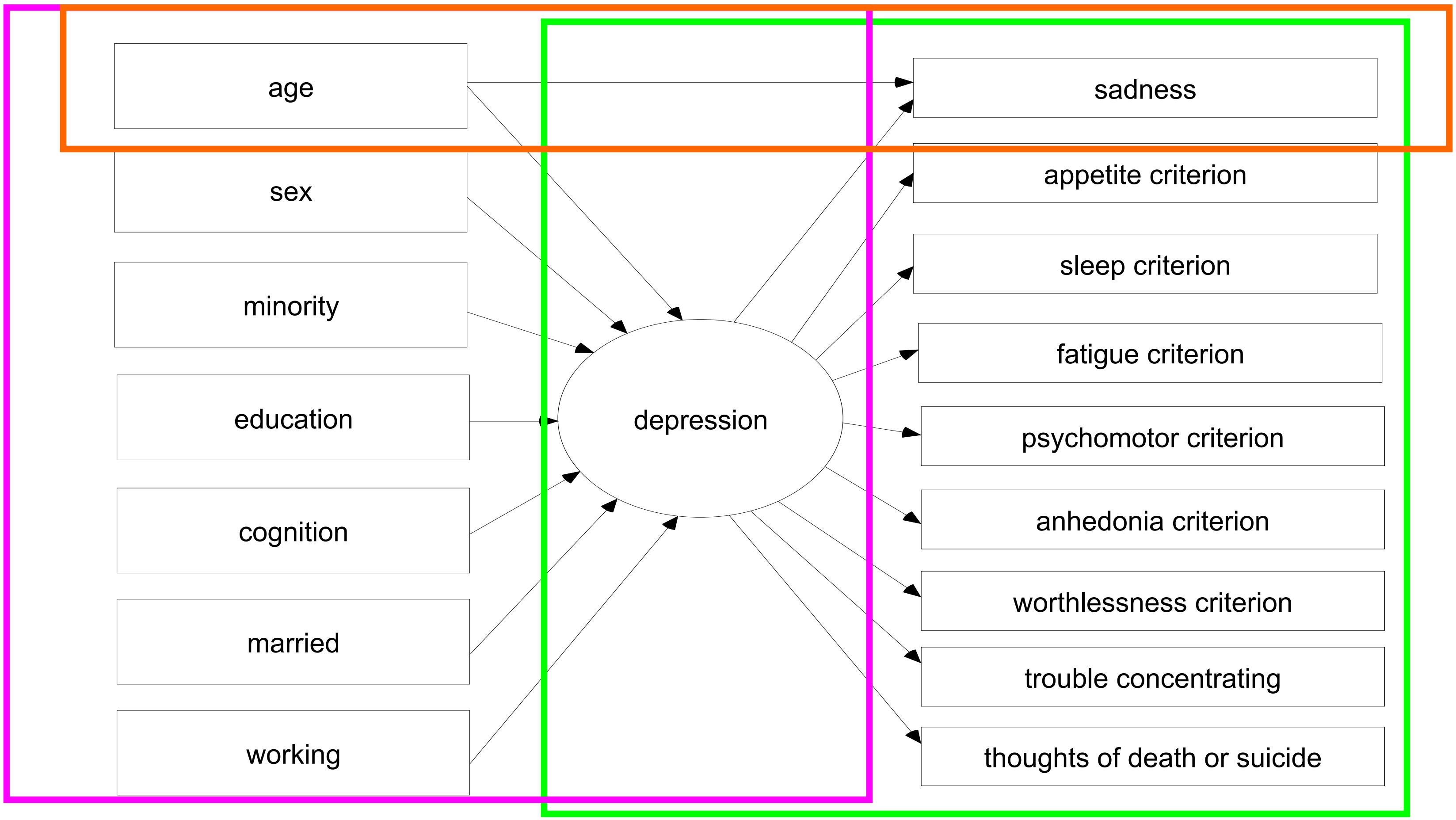
Example: Sadness in Older Adults

(Gallo, J.J., P. V. Rabins, and J. C. Anthony. 1999. *Psychological Medicine* 29:341-350)

- Background: Earlier studies reported that major depression is less prevalent among the elderly and incidence rates also decline with age
- Study Hypothesis: older adults are less likely than younger adults to report sadness
- Methods:
  - A population-based 13-year follow-up study of community-living adults in Baltimore (part of ECA)
  - A subset of 1651 adults who participated at both initial interview in 1981 and follow-up interview in 1994 and were younger than 65 years of age in 1981
  - Cross-sectional analysis at follow-up using the MIMIC model

Example: Sadness in Older Adults

(Gallo, J.J., P. V. Rabins, and J. C. Anthony. 1999. *Psychological Medicine* 29:341-350)



Example: Sadness in Older Adults

(Gallo, J.J., P. V. Rabins, and J. C. Anthony. 1999. *Psychological Medicine* 29:341-350)

	Model 1	Model 2	Model 3
Direct effect of age ≥ 65 on sadness	-0.335* (-0.643, -0.027)	-0.298 (-0.602, 0.006)	-0.317* (-0.623, -0.011)
Difference in the mean level of depression comparing age ≥ 65 to < 65	-0.361* (-0.512, -0.210)	-0.581* (-0.746, -0.416)	-0.639* (-0.808, -0.470)

Model 1: adjusting for age only

Model 2: adjusting for sex, race, education, cognitive impairment, marital status, and current working status

Model 3: Model 2 – cognitive impairment

* p-value<0.05

1b) Multitrait-Multimethod SEM

Multitrait-Multimethod SEM
(Campbell & Fiske, 1959)

- Goal: separate out true variance from variance due to measurement methods
- Method: study a common set of traits by multiple methods
- Inference on
  - Convergent validity – "the tendency for different measurement operations to converge on the same underlying trait"
  - Discriminant validity – " the ability to discriminate among different traits

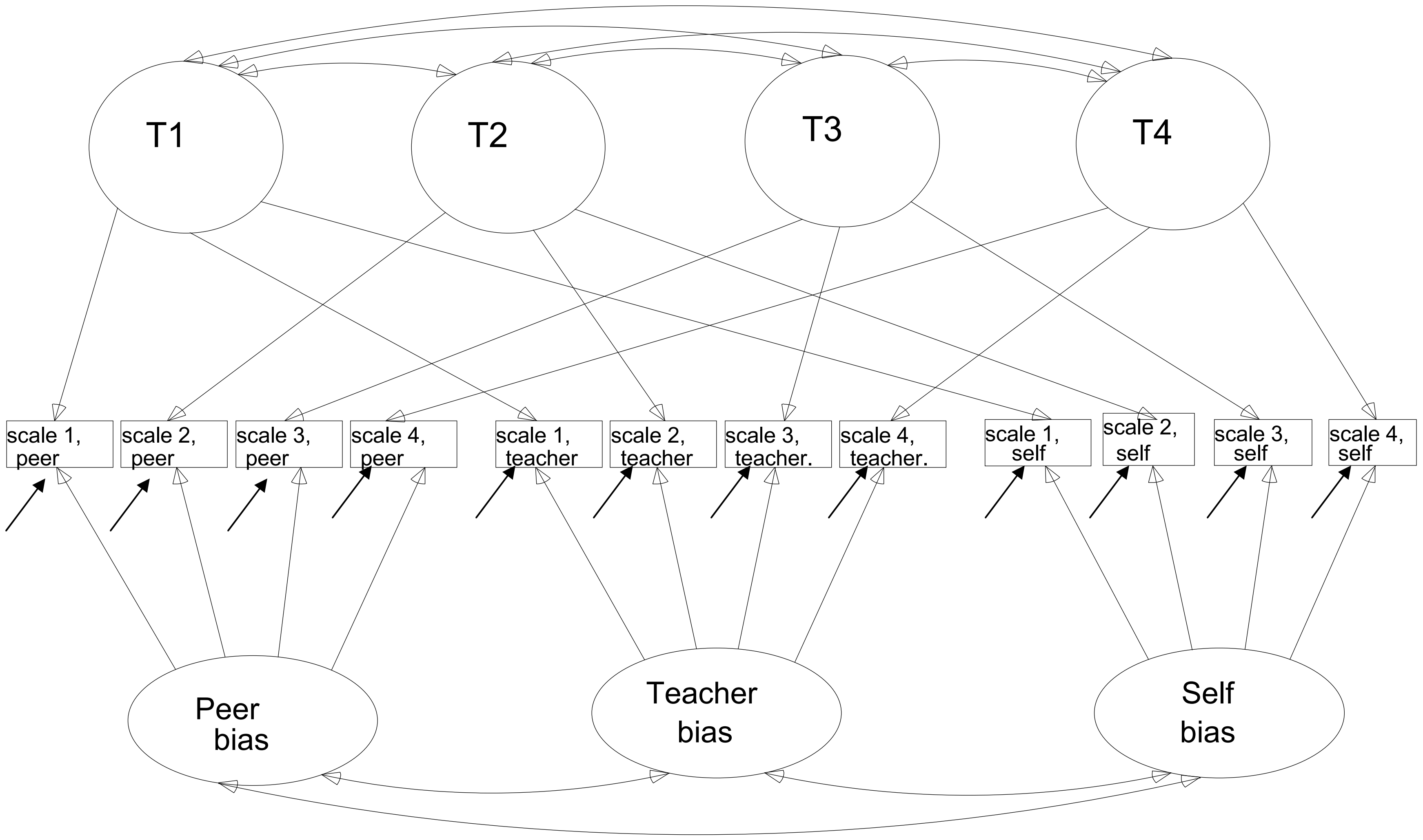
Example: Personality Traits

		Peer				Teacher				Self			
		1	2	3	4	1	2	3	4	1	2	3	4
P	1	(.rr)											
e	2	.aa	(.rr)										
e	3	.aa	.aa	(.rr)									
r	4	.aa	.aa	.aa	(.rr)								
T	1	*.cc*	.bb	.bb	.bb	(.rr)							
e	2	.bb	*.cc*	.bb	.bb	.aa	(.rr)						
a	3	.bb	.bb	*.cc*	.bb	.aa	.aa	(.rr)					
c	4	.bb	.bb	.bb	*.cc*	.aa	.aa	.aa	(.rr)				
S	1	*.cc*	.bb	.bb	.bb	*.cc*	.bb	.bb	.bb	(.rr)			
e	2	.bb	*.cc*	.bb	.bb	.bb	*.cc*	.bb	.bb	.aa	(.rr)		
l	3	.bb	.bb	*.cc*	.bb	.bb	.bb	*.cc*	.bb	.aa	.aa	(.rr)	
f	4	.bb	.bb	.bb	*.cc*	.bb	.bb	.bb	*.cc*	.aa	.aa	.aa	(.rr)

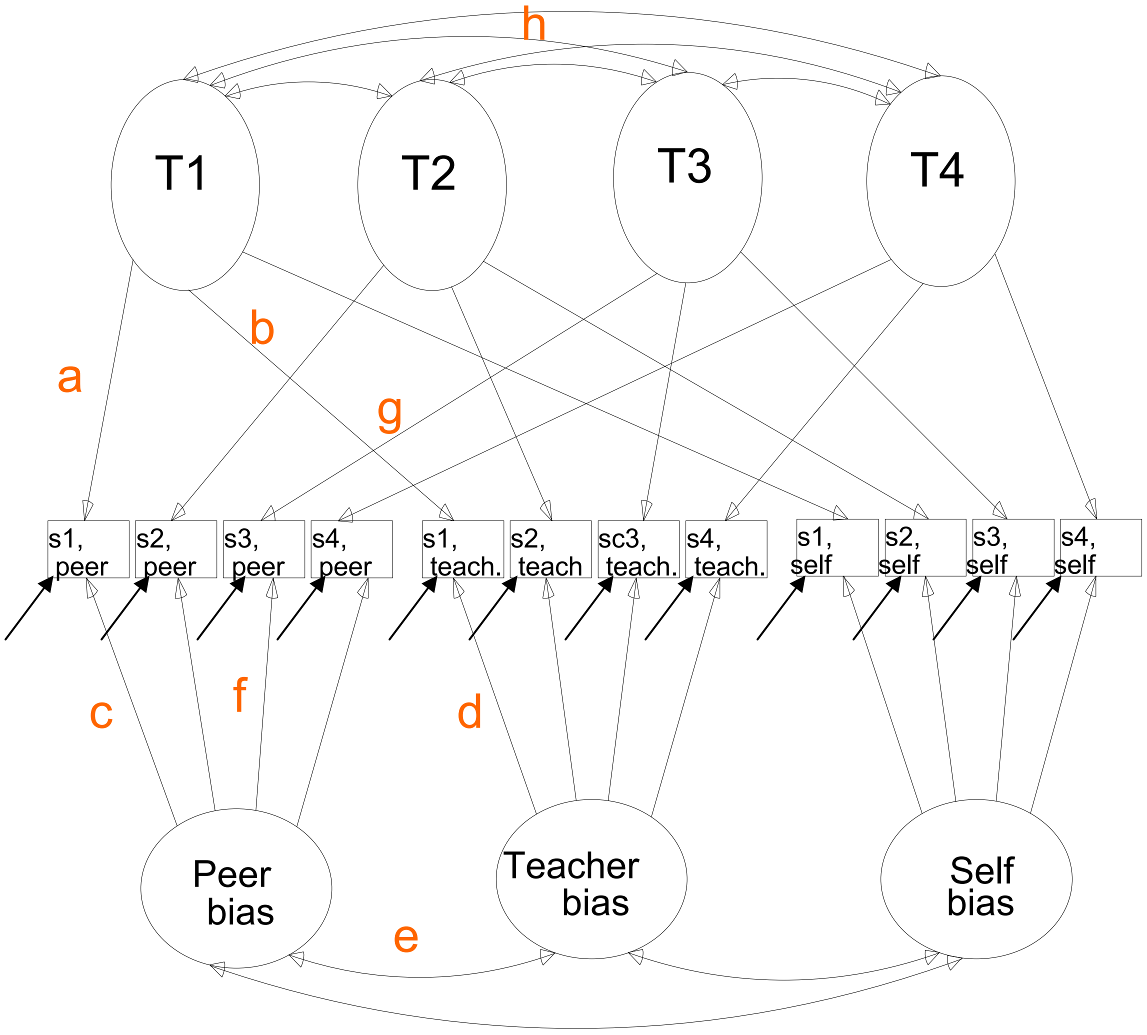
- aa – within-method cross-trait corr.
- cc – validity diagonals
- bb – cross-method cross-trait corr.
- rr = reliability coefficient

- 1 – Extraversion
- 2 – Anxiety
- 3 – impulsivity
- 4 – Academic Achievement

Example: Personality Traits



Example: Personality Traits



Within-method cross-trait Corr:

e.g. r(1p,3p)=cf+ahg

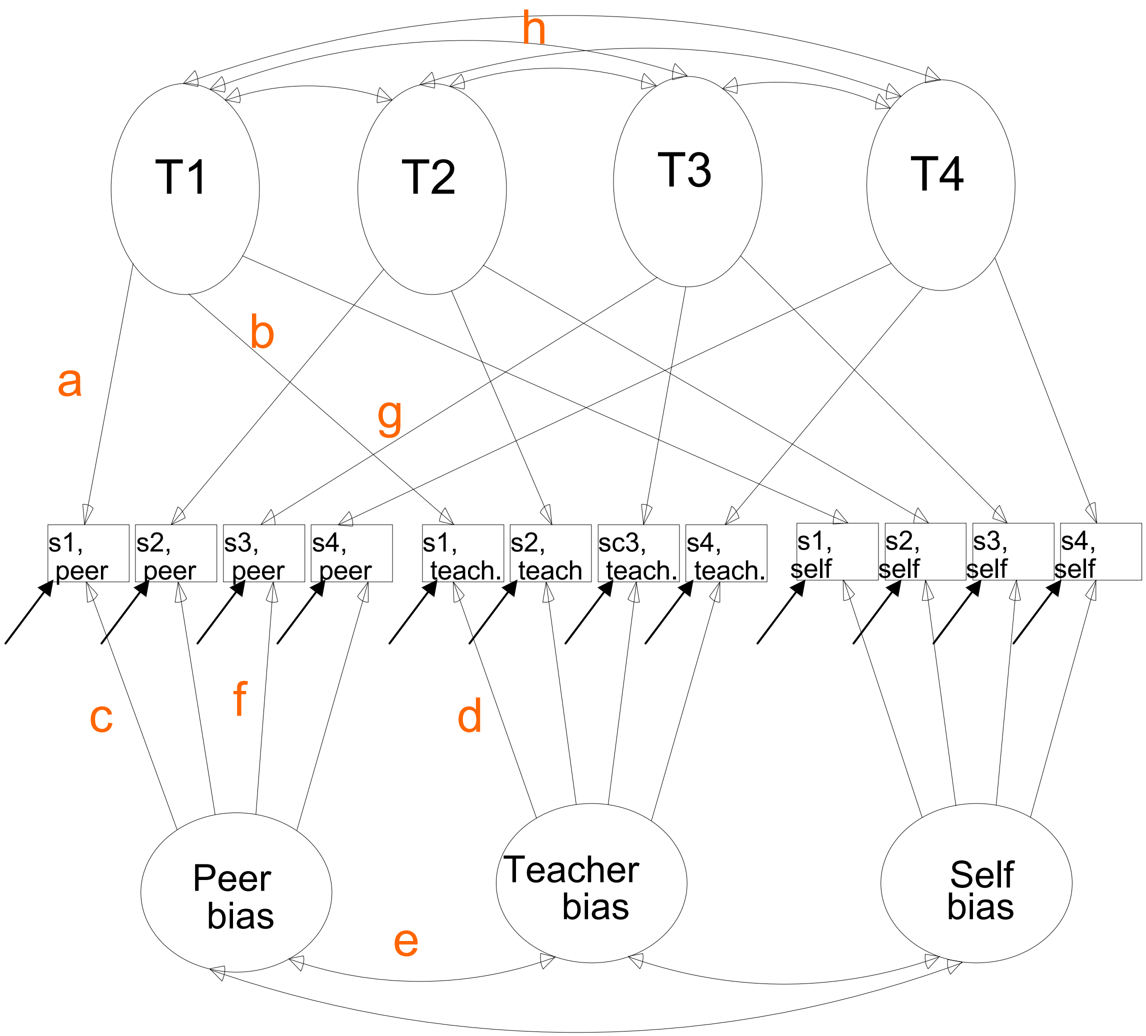
Within-trait cross-method Corr:

e.g. r(1p,1q) =ab+ced

Cross-method cross-trait Corr.

e.g. r(3p,1q)=bhg+fed

Example: Model Identification



equations = 12*13/2=78

unknowns = 45

d.f.=78-45=33 for a χ^2 goodness-of-fit test

Example: Inter-Personal Violence Scale

	Methods			
	Peer	Teacher	Self	
1. Trait Factors				
Extraversion	0.98	0.62	0.42	
Anxiety	0.77	0.91	0.35	
Impulsivity	0.78	0.64	0.42	
Motivation	0.72	0.89	0.66	
	Traits			
2. Method Factors	E	A	I	M
Peer ratings	0.15	-0.25	0.32	-0.68
Teacher ratings	0.74	-0.19	0.49	0.13
Self ratings	0.34	0.17	0.89	-0.22

(Bentler & Lee, 1979)

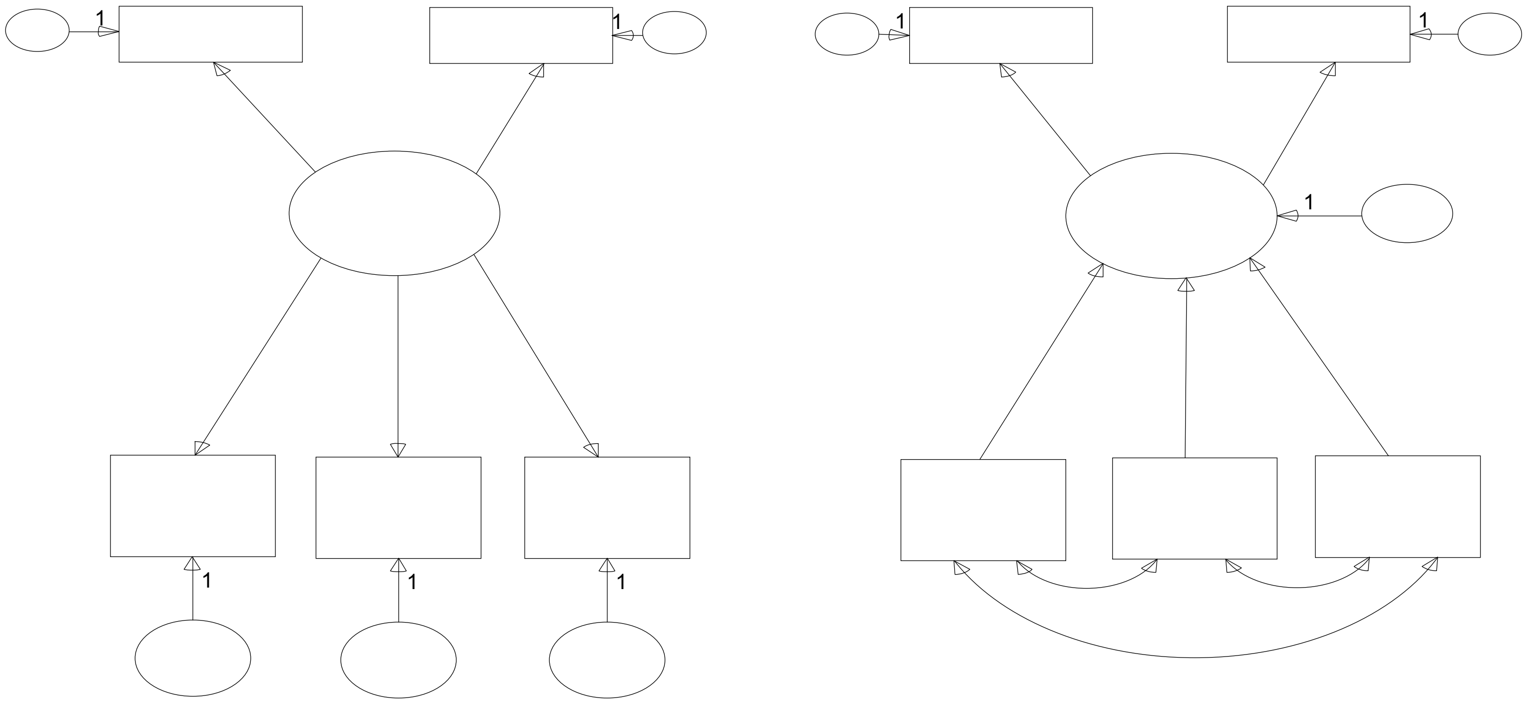
Example: Inter-Personal Violence Scale

	Methods			
	E	A	I	M
3. Trait Factor Intercorrelations				
Extraversion	1.0	-0.35	$\boxed{0.52}$	-0.24
Anxiety		1.0	-0.26	$\boxed{0.74}$
Impulsivity			1.0	-0.48
Motivation				1.0
	Traits			
2. Method Factors	P	T	S	
Peer ratings	1.0	0.08	0.04	
Teacher ratings		1.0	$\boxed{0.32}$	
Self ratings			1.0	

(Bentler & Lee, 1979)

1c) Causal vs. Effect Indicators

Cause vs. Effect Indicators



A) Effect Indicator

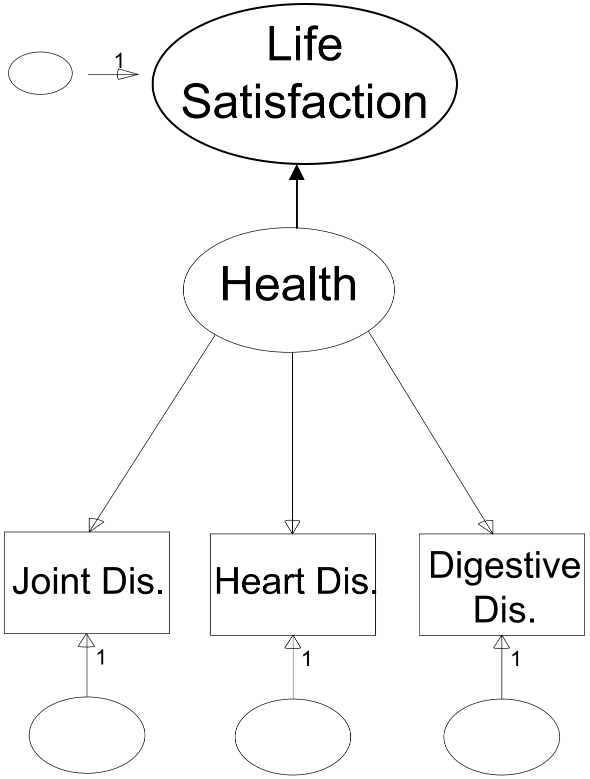
B) Causal Indicator

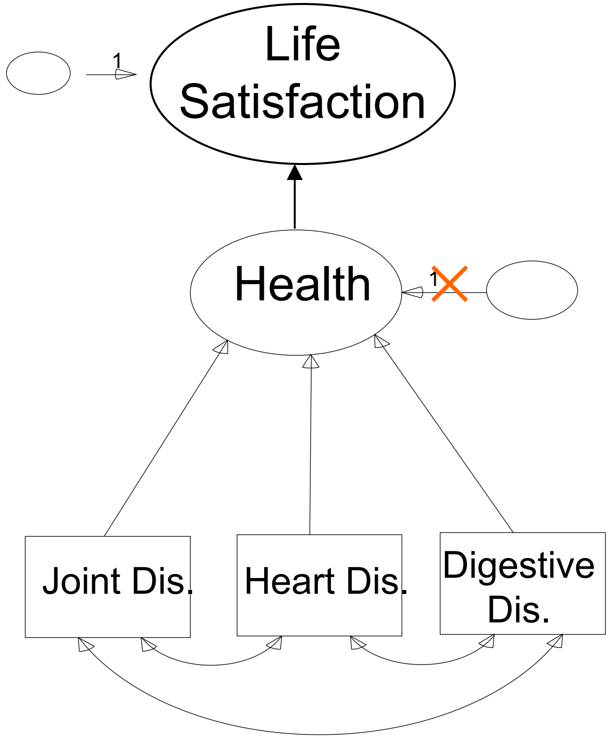
Cause vs. Effect Indicators

- Effect indicators
  - manifestations of a the latent variable (LV)
  - The LV explains the intercorrelations among the effect indicators
  - Observed correlations among the effect indicators are appropriate basis for estimating their association with the LV

- “Cause” indicators
  - Causes of the LV or
  - Relate to the LV via unmeasured common causes
  - No reason why the LV should explain the correlations among the cause indicators
  - No reason why the cause indicators should be correlated
  - Observed correlations among the cause indicators are irrelevant to their association with the LV

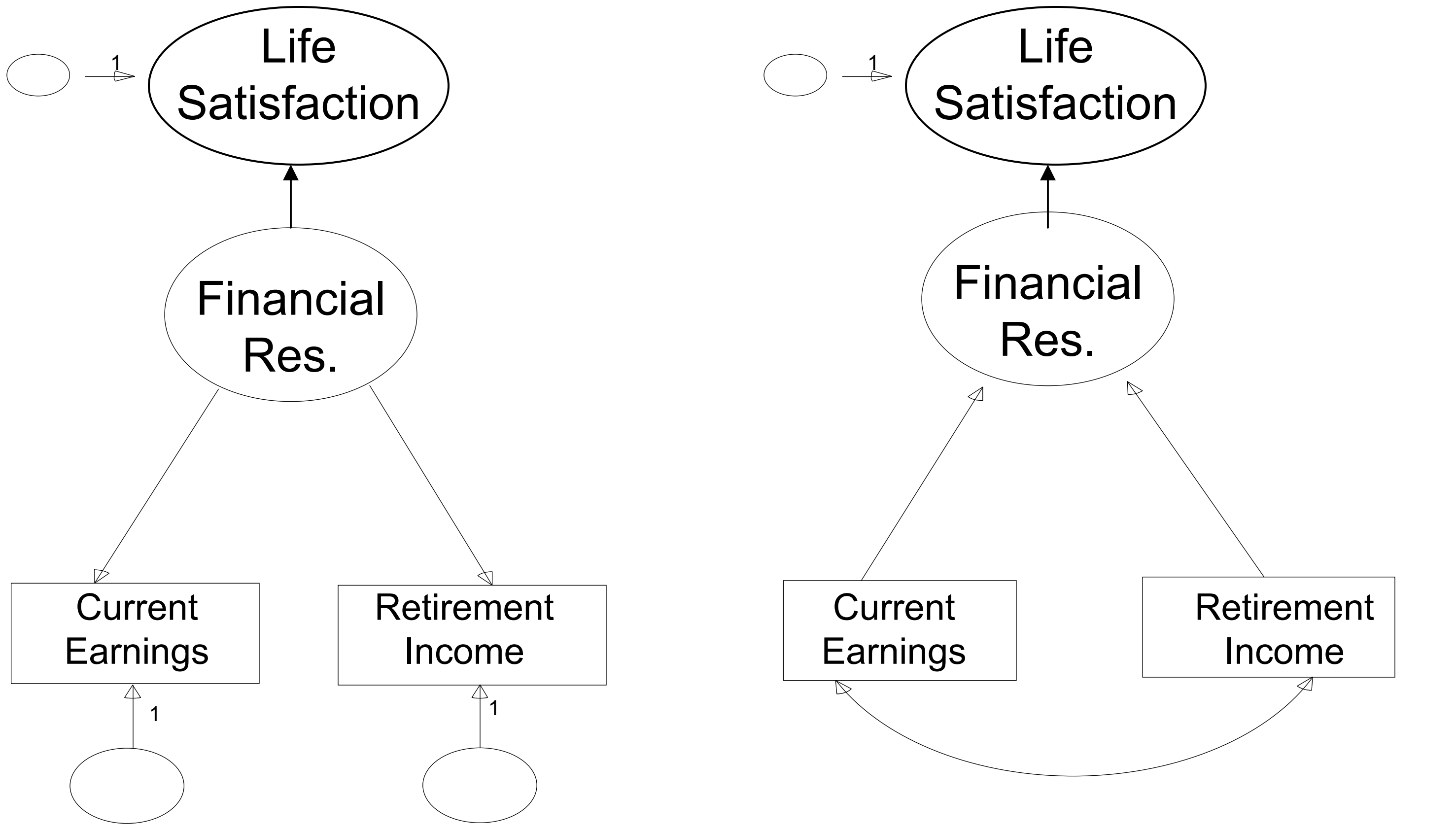
Example: Cause vs. Effect Indicators





Loehlin, pp. 219-220

Example: Cause vs. Effect Indicators



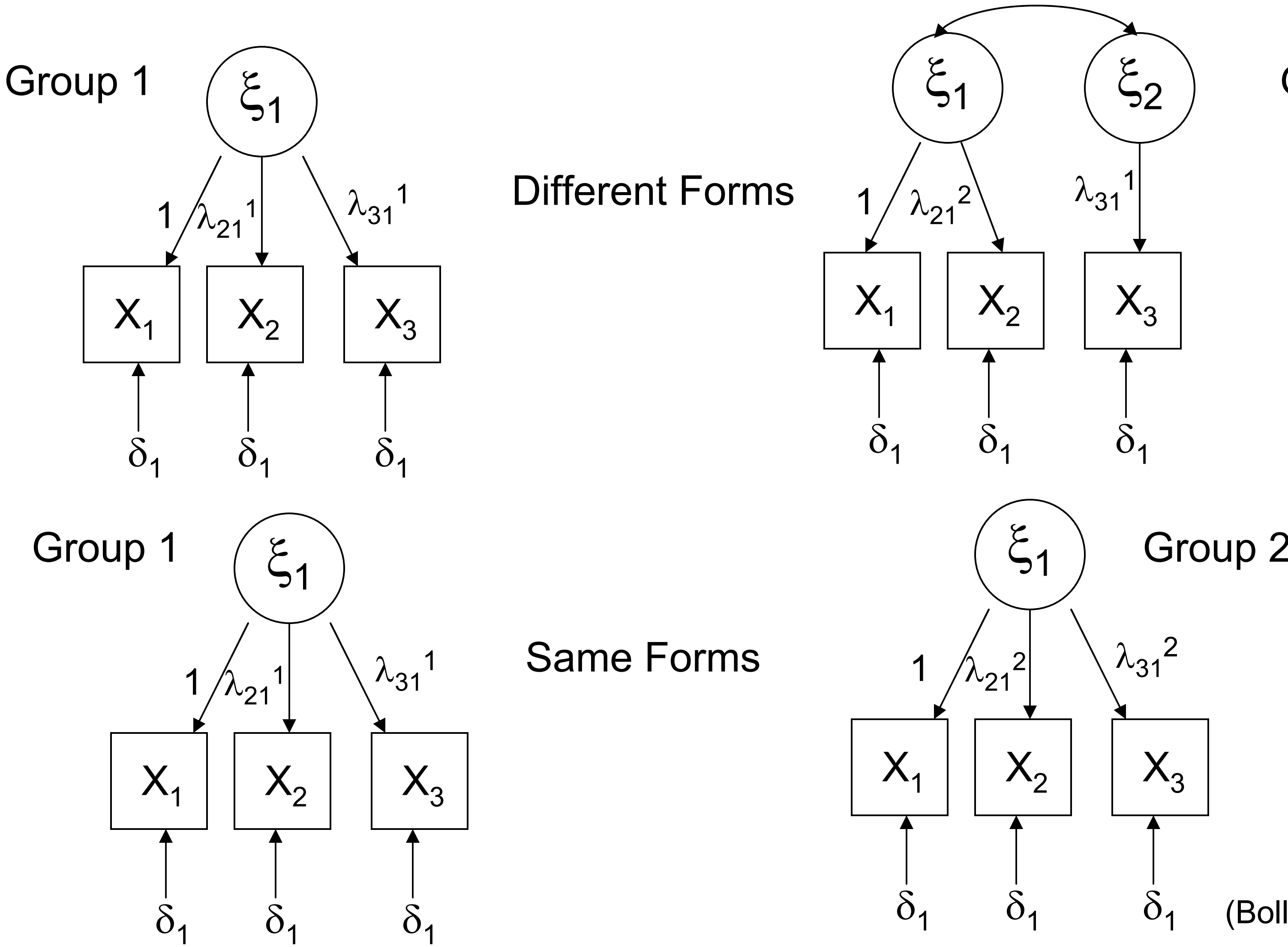
Loehlin, pp. 219-220

2. Group Comparison Models

Multiple-Group Comparisons: Example

- Suppose you had a four-indicator measure of self-esteem (exogenous) and you hypothesized that it had a different relationship with depression (endogenous) for high SES and low SES adults
- How would you test this hypothesis with regular regression analysis?
  1) You would have to *a priori* add the self-esteem indicators to make a scale.
  2) You would have to use product terms, which compound problems of reliability.
     - a) we know that the $V(t1)=\rho 1V(x1)$ and $V(t2)=\rho 2V(x2)$
     - b) so, then, the true variance of t1 times t2 (if they are not correlated) = $\rho 1\rho 2V(x1)V(x2)$

Hierarchy of Multiple-Group Comparisons



(Bollen, p. 357)

Steps in Multiple-Group Comparisons

1. Model specification for each group
2. Determine hierarchy of invariance in models
   - Similarity in model form
   - Similarity in parameter values
3. Model estimation
4. Multiple-group comparisons

Multiple-Group Comparisons: Model Fitting

- Object of analysis: S_g, i.e. group g's covariance structure, g=1, …, G
- Assumption: $\Sigma_g = \Sigma(\theta_g)$
- Fit function: $$F = \Sigma\left(\frac{N_g}{N}\right)F_g(S_g, \Sigma_g(\theta_g))$$
  i.e. a weighted sum of the fit function for each group of size N_g.
- The F_g for ML, ULS, and GLS are the same as before

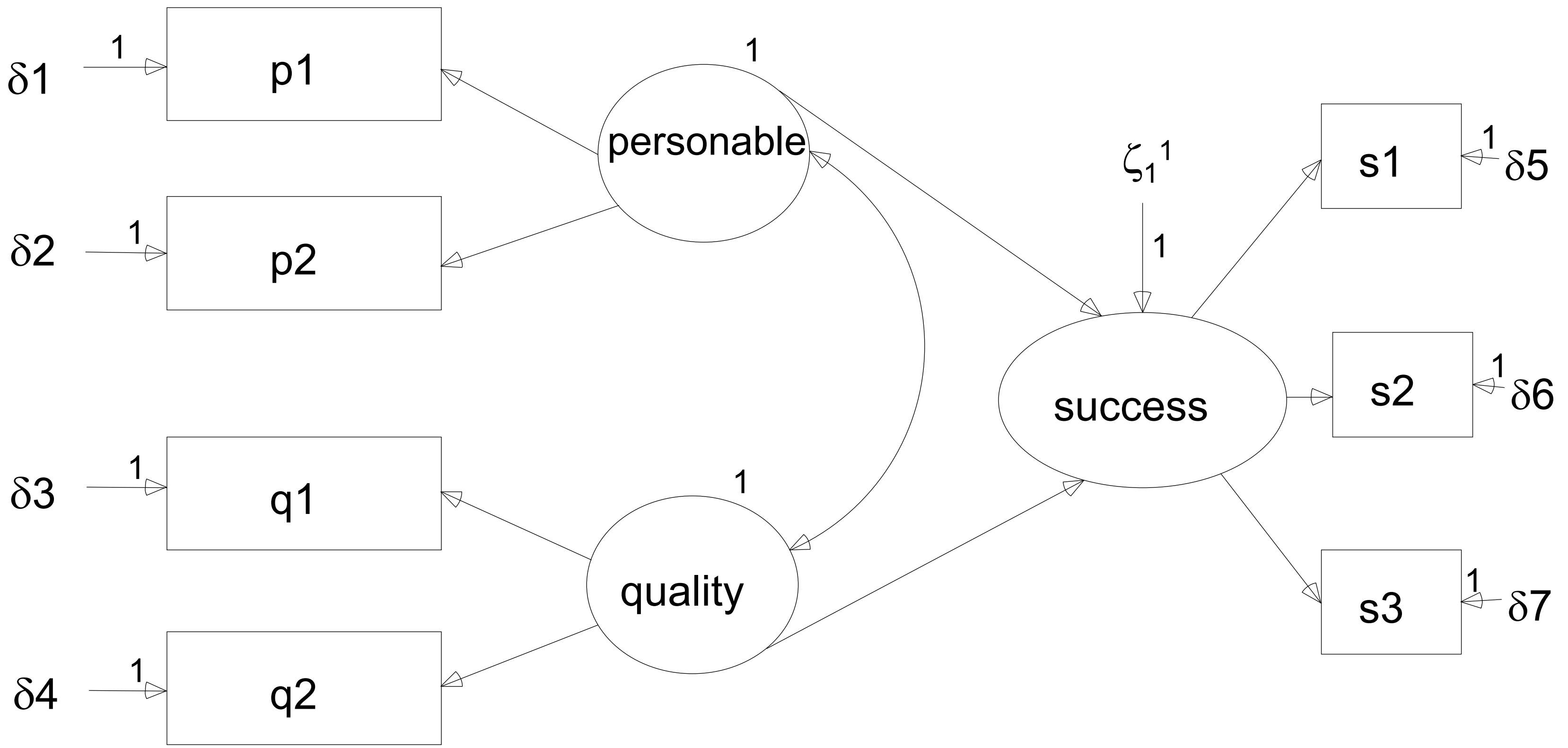
Multiple-Group Comparisons

- To assess global goodness-of-fit based on the fact: $(N-1)F \sim \chi^2$ with d.f.=$G^*[n(n+1)/2]-t$, where n is the number of observed variables, and t is the number of free parameters in <u>ALL groups</u>
- To compare nested models using the difference in χ^2 with d.f. =difference in d.f. for the two models
- Other fit statistics apply as well

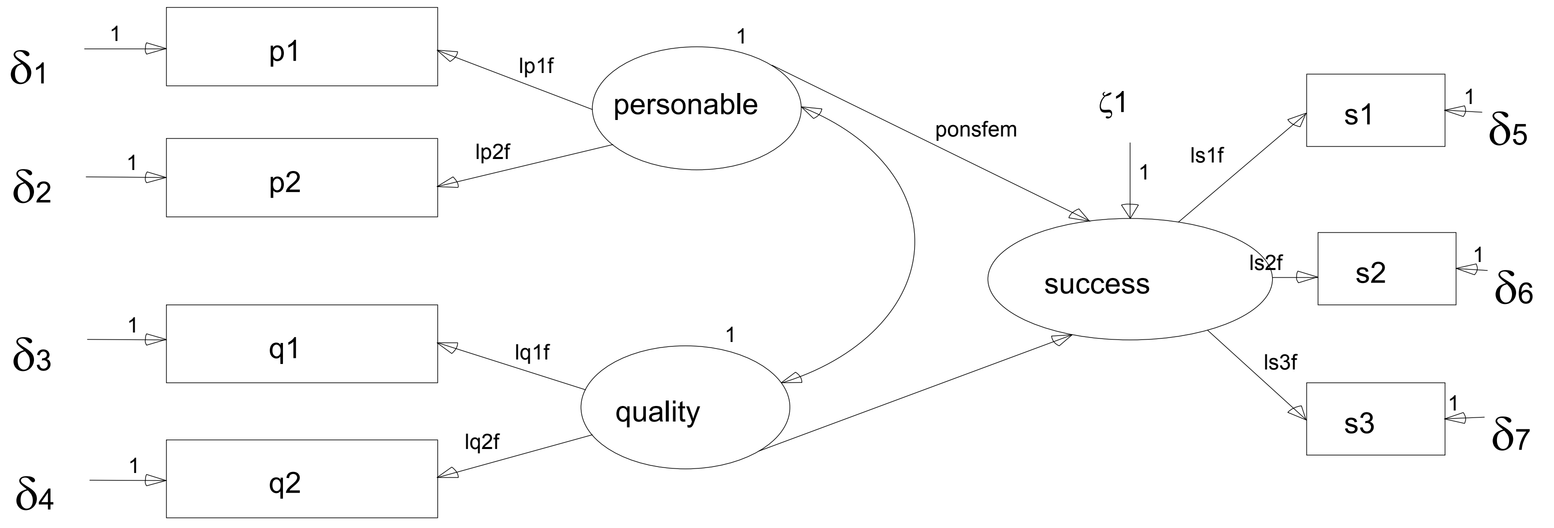
Example: Ratings of a political debate

Does gender modify the relationship between personableness and the democratic candidate's success in winning a debate?

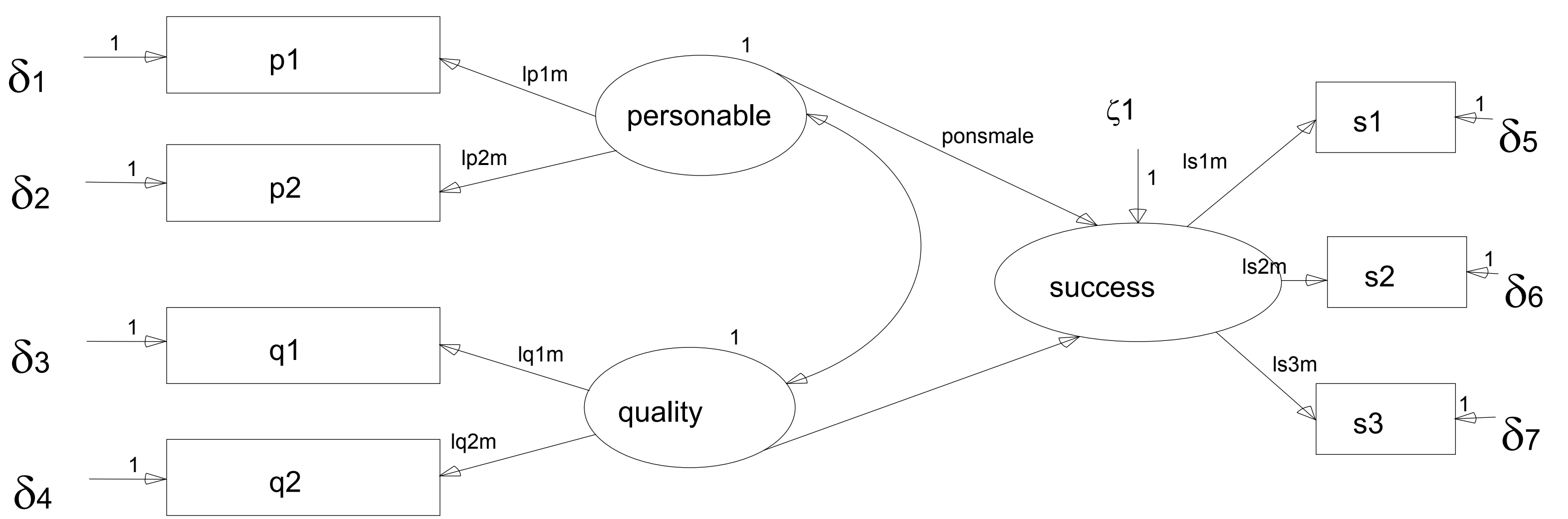
Step 1. Model specification



Women



Men



Number of parameters to estimate:

Example: Ratings of a political debate

Step 2. Hierarchy of model invariance

- Model #1: Measurement model, in which CFAs are constrained to be equal across groups (males and females)
  - lp1f=lp1m; lp2f=lp2m;
  - lq1f=lq1m; lq2f=lq2m;
  - ls1f=ls1m; ls2f=ls2m; ls3f=ls3m

- Model #2: Same as Model #1, except that the parameter of interest is constrained to be equal across groups (i.e. ponsmale=ponsfem)

- Model #3: Same as Model #1, except that the variances of residual errors of latent variable indicators are constrained to be equal across groups
  - $\Theta_\delta(m)=\Theta_\delta(f)$

Example: Ratings of a political debate

MPLUS syntax for Model #2:

```
TITLE:  Ratings of a political debate
DATA:   FILE is c:/teaching/140.658.2007/twogrp.dat;
        TYPE IS COVARIANCE;
        NOBSERVATIONS ARE 154 125;
        NGROUPS=2;
VARIABLE: NAMES ARE S1 S2 S3 P1 P2 Q1 Q2;
        USEVARIABLES ARE S1-Q2;
MODEL:
        fs BY S1* S2 S3;
        fp BY P1* P2;
        fq BY Q1* Q2;
        fs on fp (1);
        fs on fq;

        fp@1 fq@1 fs@1;
OUTPUT:
        TECH1
```

Example: Ratings of a political debate

MPLUS syntax

- Model #3:

```
MODEL:
      fs BY S1* S2 S3;
      fp BY P1* P2;
      fq BY Q1* Q2;
      fs on fp fq;
      S1 (1)
      S2 (2)
      S3 (3)
      P1 (4)
      P2 (5)
      Q1 (6)
      Q2 (7);

      fp@1 fq@1 fs@1;
OUTPUT:
      TECH1
```

Example: Ratings of a political debate

Results of two-group (males vs. females) comparisons

	DF	χ^2 (p value)	TLI	RMSEA	SRMR	BIC	Loglikelihood
Model I	29	16.9 (0.96)	1.01	0	0.02	8882.3	-4365.1
Model II	30	42.0 (0.07)	0.99	0.05	0.15	8901.7	-4377.7
Model III	36	24.1 (0.93)	1.01	0.00	0.02	8850.1	-4368.7

Model I: CFAs are constrained to be equal across groups

Model II: Model I plus the association between “peronableness” and success is constrained to be equal across groups

Model III: Model I plus the variances of residual errors of LV indicators are constrained to be equal across groups